

## DNA データバンクのゲノム解析への対処

GenBank<sup>1)</sup>, DNA Data Bank of Japan<sup>2)</sup>,  
and EMBL Data Library<sup>3)</sup>

・宮澤三造、林田秀宜  
国立遺伝学研究所  
遺伝情報研究センター  
遺伝情報分析研究室

## 要 旨

ゲノム解析の遂行には遺伝子地図、物理的地図、RFLP等多数のデータベースが必要とされるがそのなかでもDNA配列データベースは最も重要なものの一つである。ゲノム解析計画に対処するにはゲノムデータゆえの量的側面だけでなくゲノムデータという質的側面もまた考慮しなくてはならない。DNAデータバンクはこのような観点から現在以下のようなデータ収集入力システム、データ管理システムの再構築を推進している。予想されるデータの増加に対応すべく(1)著者自身によるデータエントリを支援するソフトウェアの開発。(2)データバンク間でのデータ交換、処理の自動化及びデータ受取の自動化のため、データ交換の手続き(protocol)<sup>4)</sup>を標準化。新しいDNA配列の遺伝情報に関する知見及びmap position等ゲノム解析に必須な情報を的確に表現できるよう、また他データベースとのクロスレファレンスを可能にするため(3)feature table(データの注釈テーブル)の更新<sup>5)</sup>。データの入力更新の分散処理を可能にすべく(4)現在のフラットファイルデータベースから関係データベース<sup>4)</sup>へ移行。DDBJ/EMBL/GenBankは実質的に同一のデータベースを構築しデータベースの分散管理を目指している。以上のようなDNAデータベースの再編を中心にゲノム解析計画に必須なデータベースについて議論する。

## 文献

- 1) Burks, C. et al.: The Interface between Computational Science and Nucleic Acid Sequencing, Santa Fe Institute Studies in the Sciences of Complexity, Eds. G. Bell and T. Marr(Reading, MA: Addison-Wesley), vol. VIII, 1989
- 2) Miyazawa, S.: 同上
- 3) Cameron, G. N.: Nucl. Acids Res.: 16, 1865-1867, 1988
- 4) Cinkosky, M. J., Nelson, Debra & Marr, T. G.: A technical Overview of the GenBank/HGIR Database (Los Alamos, NM: GenBank/HGIR), 1989
- 5) DDBJ, EMBL Data Library & GenBank: The DDBJ/EMBL/GenBank Feature Table: Definition, version 1, 1988